How does weather explain New York shooting's rates?

I apply Bayesian Multilevel to model the rates of shooting in the City of Greater New York. While the second level of the model pools census information across space, the higher level simulate the trend and variability in response to the daily weather. This arrangement allows us to explore the variability of response with regard to the different social/economic background. If time permits, at the later stages of this investigation, I incorporate the effect of air quality (using AQI) to inform the rates of gun violence.

Preliminary Analysis : Temporal Varaiblity

For this analysis, I use multiple sources of information. This includes NYPD historical rate dataset (2006-2018), NCDC daily weather data, NYC census data. Later, I will include the AQI (Air Quality Index) data to incorporate the relevancy of air pollution to the occurrence of shooting. The time-span of observation considered in this analysis is 2006-2018.

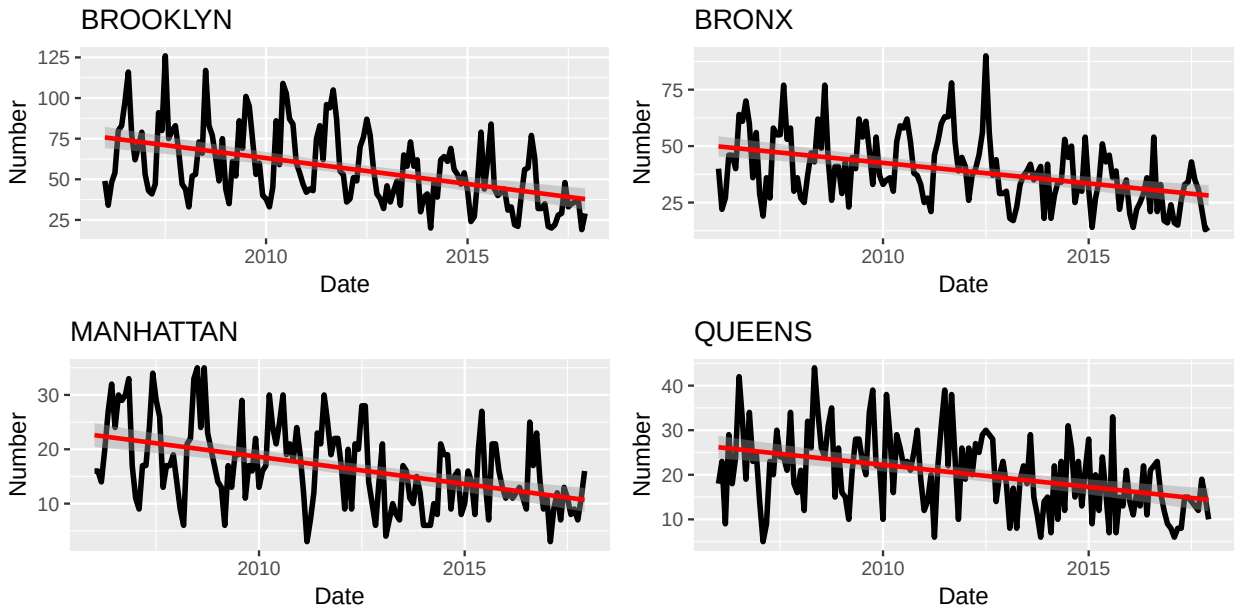
A sample of NYDP shooting dataset looks like the following:

```
head(sample_of_data,4)
```

```
##       Borough Precinct Murder Age Group Sex      Lat       Lon      Dates
## 31   BROOKLYN       67  FALSE                40.63888 -73.94394 2011-09-09
## 27     QUEENS      104  FALSE     18-24    M 40.70444 -73.91197 2011-09-09
## 90     QUEENS      105  FALSE     25-44    M 40.66801 -73.74069 2014-09-08
## 99  MANHATTAN       30  FALSE     18-24    M 40.82747 -73.94597 2012-09-08
```

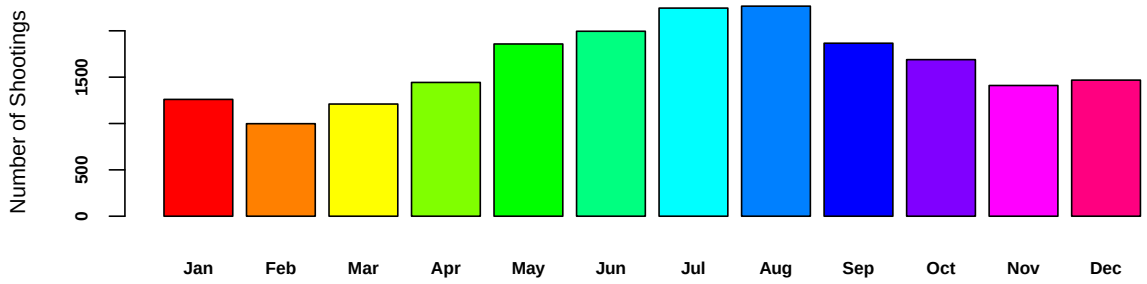
During recent years, the New York shooting rates constantly decreased. This negative trends in year-to-year data can root from diffrent sources (such as increase in the number of police stations, education and etc.)

```
grid.arrange(p1, p2, p3, p4, ncol = 2)
```

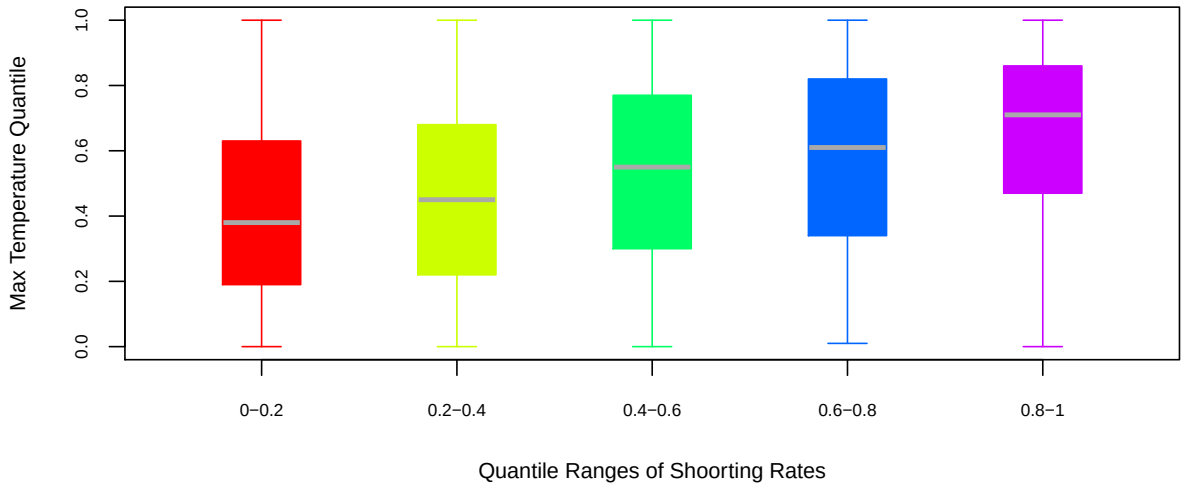


However, the shooting rates present a seasonal pattern, suggesting a correlation with weather variability. The Following figure shows that the number of incidents in the warmer months is higher than the shooting rates in colder seasons.

```
df_plot <- aggregate(STATISTICAL_MURDER_FLAG~MONTH,as.data.frame(df),length)
barplot(sapply(1:12,function(x) length(df$STATISTICAL_MURDER_FLAG[which(df$MONTH==x)])),
        col=rainbow(12),names=month.abb[1:12],ylab='Number of Shootings',
        font=2,cex.axis = 0.7,cex.names=0.7,cex.lab=0.8)
```

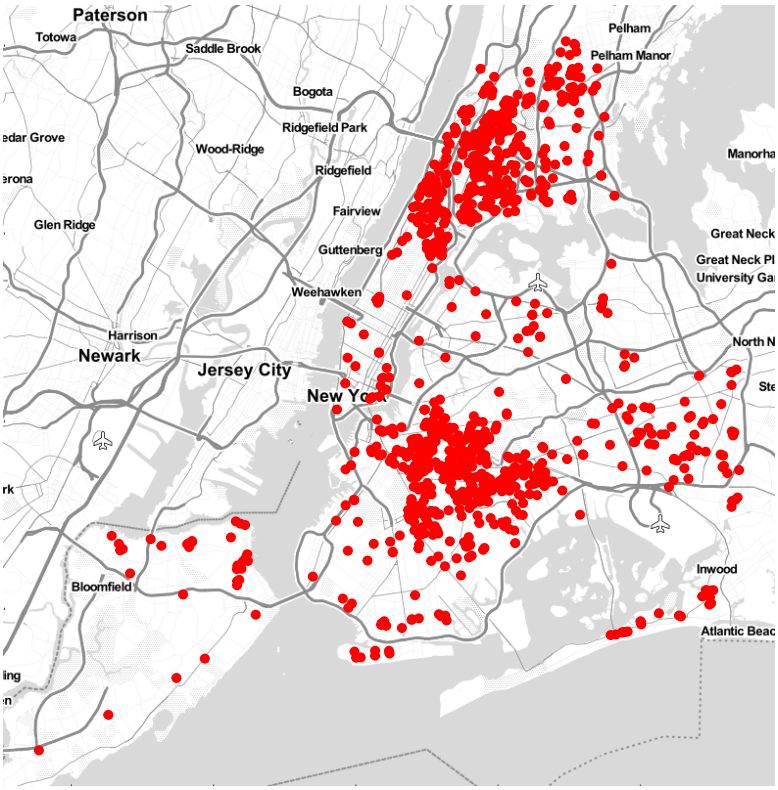


The additional evidence is found in comparing the daily shooting incidents with weather variables. On the daily number of incidents, I extract 6 different classes with different range of quantiles. Further, I compare the outcome with the quantile of the corresponding maximum daily temperature. The boxplots suggest that the higher number of shooting correlates with higher temperature.

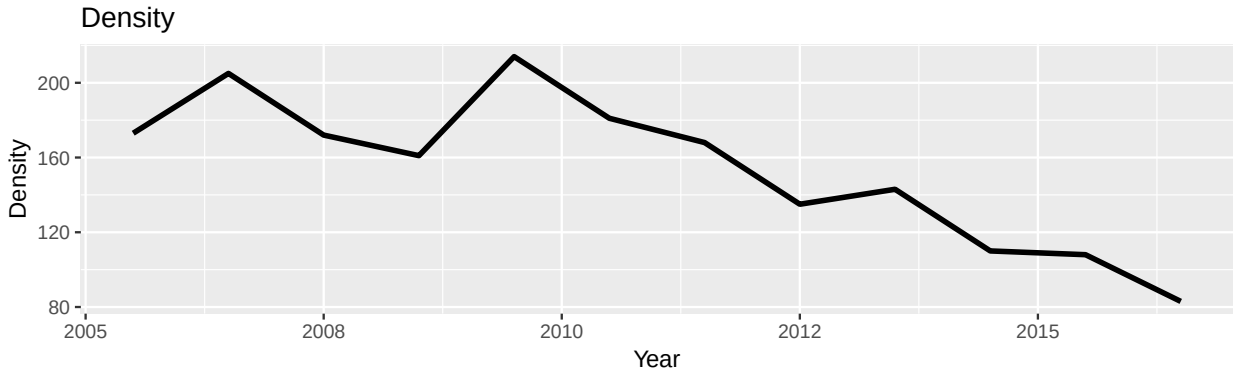


Preliminary Analysis : Spatial Varaiblity

The spatial distribution of shooting incident yields useable information on how the affected areas changed over time. The following map demonstrates the location of gun violence in 2007.



Using the Kernel estimators, I looked at the density of the number of incindents within 2-km radius. We observe that the density of crimes decrease over time.

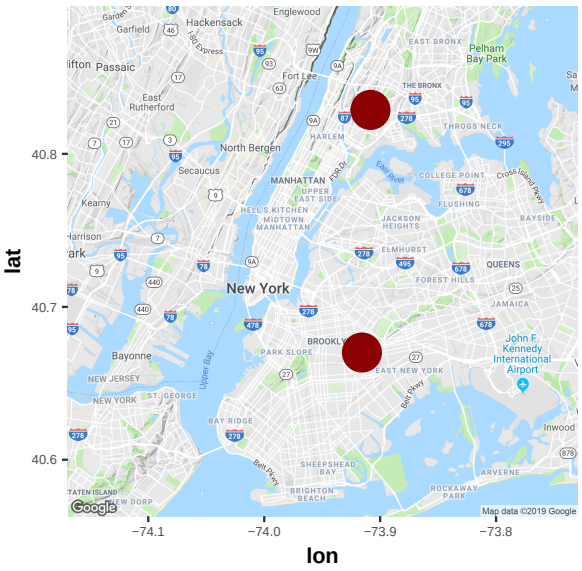


Moreover, the following maps indicate the highest density points throughout the years. We observe that the highest density point of crimes from Brooklyn in 2009 shifts to the Bronx in 2014. In this project, I study this shift and relate that to the social/economic background of the neighborhood. I ask as to what extent the change is latter, modulate the spatial shift in the crime rates.

2008

2011

2014



2017